PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 : C07K 14/00

A2

(11) International Publication Number: WO 00/37492

(43) International Publication Date: 29 June 2000 (29.06.00)

(21) International Application Number: PCT/US99/30932

(22) International Filing Date: 22 December 1999 (22.12.99)

(30) Priority Data:
60/113,344 22 December 1998 (22.12.98) US

(71) Applicant *(for all designated States except US)*: ELI LILLY AND COMPANY [US/US]; Lilly Corporate Center, Indianapolis, IN 46285 (US).

(72) Inventor; and

(75) Inventor/Applicant *(for US only)*: NA, Songqing [CN/US]; 10756 Independence Way, Carmel, IN 46032 (US).

(74) Agents: PLANT, Thomas, G. et al.; Eli Lilly and Company, Lilly Corporate Center, Indianapolis, IN 46285 (US).

(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published

Without international search report and to be republished upon receipt of that report.

(54) Title: huHDGFh NUCLEIC ACIDS, POLYPEPTIDES, METHODS AND USES THEREOF

(57) Abstract

The present invention relates to at least one novel huHDGFh polypeptide, including isolated nucleic acids that encode at least one huHDGFh polypeptide, huHDGFh polypeptides, vectors, host cells, transgenics, chimerics, and methods of making and using thereof, as well as huHDGFh–specific antibodies and methods.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

huHDGFh NUCLEIC ACIDS, POLYPEPTIDES, METHODS AND USES THEREOF

This application claims priority of U.S. Provisional Application Serial No. 60/113,344, filed December 22, 1998.

The present invention relates to compounds and compositions comprising novel human homologs of hepatoma-derived growth factor homologous (huHDGFh) polypeptides, nucleic acids, host cells, transgenics, chimerics, antibodies, compositions, and methods of making and using thereof.

Cell growth is regulated by various growth factors and cytokines, which bind to specific membrane receptors to trigger a cascade of intracellular biochemical signals to the activation of transcription factors, resulting in the activation and repression of various subsets of genes (Aaronson, S.A., *Science*, 254:1146-1153 (1991)).

Hepatoma-derived growth factor (HDGF) is a heparin-binding protein which is mitogenic for fibroblasts and some hepatoma cells (Nakamura, H. *et al.*, *J. Biol. Chem.*, 269(40):25143-25149 (1994); Nakamura, H. *et al.*, *Clin. Chim. Acta* 183, 273-284 (1989)). HDGF was purified from the conditioned medium of a human hepatoma-derived cell line,

HuH-7 by tritiated thymidine incorporation into Swiss 3T3 cells. HDGF has no signal peptide, yet is secreted into the medium of COS-7 cells after transfection of the cDNA clone. It is a heparin-binding protein and is ubiquitously expressed in several tumor-derived cell lines and tissues. It is localized in the cytoplasm of hepatoma cells and has strong growth stimulating activity. Such cytokines or growth factor like molecules related to HDGF satisfy a need in the art by providing new diagnostic or therapeutic compositions useful in diagnosing and treating infections; autoimmune disorders, vascular diseases and cancers.

Accordingly, there is a need to provide human homologs of HDGF polypeptides, nucleic acids, host cells, transgenics, chimerics, as well as methods of making and using thereof.

The present invention provides isolated nucleic acids and encoded huHDGFh polypeptides, including specified fragments and variants thereof, as well as huHDGFh compositions, probes, primers, vectors, host cells, antibodies, transgenics, chimerics and methods of making and using thereof, as described and enabled herein.

The present invention provides, in one aspect, isolated nucleic acid molecules comprising or complementary to a polynucleotide encoding specific huHDGFh polypeptides, as fragments or specified variants, comprising at least one domain thereof.

Such polypeptides are provided as non-limiting examples by the corresponding domains, specified fragments and/or

specified variants of huHDGFh polypeptides corresponding to at least 90-100% of SEQ ID NO:2.

The present invention further provides recombinant vectors, comprising 1-40 of said isolated huHDGFh nucleic acid molecules of the present invention, host cells containing such nucleic acids and/or recombinant vectors, as well as methods of making and/or using such nucleic acid, vectors and/or host cells.

The present invention also provides methods of making or using such nucleic acids, vectors and/or host cells, such as but not limited to, using them for the production of huHDGFh nucleic acids and/or polypeptides by known recombinant, synthetic and/or purification techniques, based on the teaching and guidance presented herein in combination with what is known in the art.

The present invention also provides an isolated huHDGFh polypeptide, comprising at least one fragment, domain or specified variant of at least 90-100% of the contiguous amino acids of at least one portion of SEQ ID NO:2.

The present invention also provides an isolated huHDGFh polypeptide as described herein, wherein the polypeptide further comprises at least one specified substitution, insertion or deletion corresponding to portions or residues of SEQ ID NO:2.

The present invention also provides an isolated huHDGFh polypeptide as described herein, wherein the polypeptide has at least one activity, such as, but not limited to, inducing cell proliferation, hematopoiesis, lymphocyte proliferation, and angiogenesis (Nakamura, H. *et al., Clin. Chim. Acta*

183:273-284 (1989); Varnum-Finney B. *et al, Blood*, 91:4084-4091 (1998); McKenzie, G. *et al, Current Biology*. 8(6):339-42 (1998); Cid, M. *et al., J. Clinical Invest.*, 91:977-985 (1993); Schnaper, W., *et al., J. Cellular Physiol.*, 165:107-118 (1995)). A huHDGFh polypeptide can thus be screened for a corresponding activity according to known methods.

The present invention also provides a composition comprising an isolated huHDGFh nucleic acid and/or polypeptide as described herein and a carrier or diluent. The carrier or diluent can optionally be pharmaceutically acceptable, according to known methods.

The present invention also provides an isolated nucleic acid probe, primer or fragment, as described herein, wherein the nucleic acid comprises a polynucleotide of at least 10 nucleotides, corresponding or complementary to at least 10 nucleotides of SEQ ID NO:1.

The present invention also provides a recombinant vector comprising an isolated huHDGFh nucleic acid as described herein.

The present invention also provides a host cell, comprising an isolated huHDGFh nucleic acid as described herein.

The present invention also provides a method for constructing a recombinant host cell that expresses a huHDGFh polypeptide, comprising introducing into the host cell a huHDGFh nucleic acid in replicatable form as described herein to provide the recombinant host cell. The present invention also provides a recombinant host cell provided by a method as described herein.

The present invention also provides a method for expressing at least one huHDGFh polypeptide in a recombinant host cell, comprising culturing a recombinant host cell as described herein under conditions wherein at least one huHDGFh polypeptide is expressed in detectable or recoverable amounts.

The present invention also provides an isolated huHDGFh polypeptide produced by a recombinant, synthetic, and/or any suitable purification method as described herein and/or as known in the art.

The present invention also provides a huHDGFh antibody or fragment, comprising a polyclonal and/or monoclonal antibody or fragment that specifically binds at least one epitope specific to at least one huHDGFh polypeptide as described herein.

The present invention also provides a method for producing a huHDGFh antibody or antibody fragment, comprising generating the antibody or fragment that binds at least one epitope that is specific to an isolated huHDGFh polypeptide as described herein, the generating done by knowing recombinant, synthetic and/or hybridoma methods.

The present invention also provides a huHDGFh antibody or fragment produced by a method as described herein or as known in the art.

The present invention also provides a method for identifying compounds that bind a huHDGFh polypeptide, comprising

a) admixing at least one isolated huHDGFh polypeptide as described herein with a test compound or composition; and

b) detecting at least one binding interaction between the polypeptide and the compound or composition, optionally further comprising detecting a change in biological activity, such as a reduction or increase.

The present invention provides isolated, recombinant and/or synthetic nucleic acid molecules comprising at least one polynucleotide encoding at least one huHDGFh polypeptide comprising specific full length sequences, fragments and specified variants thereof, such polypeptides, and methods of making and using said nucleic acids and polypeptides thereof. A huHDGFh polypeptide of the invention comprises at least one fragment, domain, and/or specified variant as a portion or fragment of a huHDGFh protein as described herein.

Utility

The present invention also provides at least one utility by providing isolated nucleic acid comprising polynucleotides of sufficient length and complementarity to a huHDGFh nucleic acid for use as probes or amplification primers in the detection, quantitation, or isolation of gene sequences or transcripts. For example, isolated nucleic acids of the present invention can be used as probes for detecting deficiencies in the level of mRNA, in screens for detection of mutations in at least one huHDGFh gene (e.g., substitutions, deletions, or additions), or for monitoring upregulation of expression of said gene, or changes in biological activity as described herein in screening assays

of compounds, and/or for detection of any number of allelic variants (polymorphisms or isoforms) of the gene.

The isolated nucleic acids of the present invention can also be used for recombinant expression of huHDGFh polypeptides, or for use as immunogens in the preparation and/or screening of antibodies. The isolated nucleic acids of the present invention can also be employed for use in sense or antisense suppression of one or more huHDGFh genes or nucleic acids, in a host cell, or tissue *in vivo* or *in vitro*. Attachment of chemical agents which bind, intercalate, cleave and/or crosslink to the isolated nucleic acids of the present invention can also be used to modulate transcription or translation of at least one nucleic acid disclosed herein.

Citations

All publications or patents cited herein are entirely incorporated herein by reference as they show the state of the art at the time of the present invention to provide description and enablement of the present invention. Publications refer to scientific, patent publication or any other information available in any media format, including all recorded, electronic or printed formats. The following citations are entirely incorporated by reference: Ausubel, *et al.*, ed., *Current Protocols in Molecular Biology*, Greene Publishing, NY, NY (1987-1998); Sambrook, *et al.*, *Molecular Cloning:* A Laboratory Manual, 2nd Edition, Cold Spring Harbor, NY (1989); Harlow and Lane, Antibodies, a Laboratory Manual, Cold Spring Harbor, NY (1989); Colligan, *et al.*,

eds., *Current Protocols in Immunology*, Greene Publishing, NY (1994-1998).

Definitions

The following definitions of terms are intended to correspond to those as well known in the art. The following terms are therefore not limited to the definitions given, but are used according to the state of the art, as demonstrated by cited and/or contemporary publications or patents.

A "polynucleotide" comprises at least 10-20 nucleotides of a nucleic acid (RNA, DNA or combination thereof), provided by any means, such as synthetic, recombinant isolation or purification method steps.

The term "antibody" refers to intact molecules as well as to fragments thereof, such as Fa, $F(ab')_2$, and Fv fragments which are capable of binding the eptitopic determinant. Antibodies that bind huHDGFh polypeptides can be prepared using intac polypeptides or using fragments containing small peptides of interest as the immunizing antigen. The polypeptide or oligopeptide used to immunize an animal (e.g., a mouse, a rat, or a rabbit) can be derived from the translation of RNA, or synthesized chemically, and can be conjugated to a carrier protein if desired.

The terms "complementary" or "complementarity" as used herein refer to the capacity of purine, pyrimidine, synthetic or modified nucleotides to associate by partial or complete complementarity through hydrogen or other bonding to form partial or complete double- or triple-stranded nucleic acid molecules. The following base pairs occur by

complete complementarity: (i) guanine (G) and cytosine (C); (ii) adenine (A) and thymine (T); and adenine (A) and uracil (U). "Partial complementarity" refers to association of two or more bases by one or more hydrogen bonds or attraction that is less than the complete complementarity as described above. Partial or complete complementarity can occur between any two nucleotides, including naturally occurring or modified bases, e.g., as listed in 37 CFR § 1.822. All such nucleotides are included in polynucleotides of the invention as described herein.

The term "conservative" in reference to an amino acid change or substitution is intended to indicate an amino acid has been replaced with a similar amino acid. Similar amino acids are amino acids that, because of size, charge, polarity and conformation, are more readily substituted without significantly affecting the structure and/or function of the protein. Thus, one skilled in the art generally does not expect a "conservative" amino acid change or substitution to result in any measurable difference in any particular characteristic, property, and/or activity of a polypeptide having a particular conservative amino acid substitution. Specific examples of amino acid changes or substitutions considered to be conservative are known in the art. These examples include, but are not limited to, the non-polar amino acids Glycine, Alanine, Valine, Isoleucine, and Leucine; the non-polar aromatic amino acids Phenylalanine, Tryptophan, and Tyrosine; the neutral polar amino acids Serine, Threonine, Cysteine, Glutamine, Asparagine, and Methionine; the negatively charged amino acids Lysine, Arginine, and Histidine; the positively

charged amino acids Aspartate and Glutamate, represent groups of conservative amino acids. Substitution of any one for another in the same group would generally be considered to be a "conservative" substitution by one skilled in the art (See generally, James D. Watson et al., *Molecular Biology of the Gene* (1987)).

The term "fusion protein" denotes a hybrid protein molecule not found in nature comprising a translational fusion or enzymatic fusion in which two or more different proteins or fragments thereof are covalently linked on a single polypeptide chain. The term "polypeptide" also includes such fusion proteins.

"Host cell" refers to any eucaryotic, procaryotic, or fusion or other cell or pseudo cell or membrane-containing construct that is suitable for propagating and/or expressing an isolated nucleic acid that is introduced into a host cell by any suitable means known in the art (e.g., but not limited to, transformation or transfection, or the like), or induced to express an endogenous nucleic acid encoding a huHDGFh polypeptide according to the present invention. The cell can be part of a tissue or organism, isolated in culture or in any other suitable form.

The term "hybridization" as used herein refers to a process in which a partially or completely single-stranded nucleic acid molecule joins with a complementary strand through nucleotide base pairing. Hybridization can occur under conditions of low, moderate or high stringency, with high stringency preferred. The degree of hybridization depends upon, for example, the degree of homology, the

stringency conditions, and the length of hybridizing strands as known in the art.

By "isolated" nucleic acid molecule(s) is intended a nucleic acid molecule, DNA, RNA, or both which has been removed from its native or naturally occurring environment. For example, recombinant nucleic acid molecules contained or generated in culture, a vector and/or a host cell are considered isolated for the purposes of the present invention. Further examples of isolated nucleic acid molecules include recombinant nucleic acid molecules maintained in heterologous host cells or purified (partially or substantially) nucleic acid molecules in solution. Isolated RNA molecules include in vivo or in vitro RNA transcripts of the nucleic acid molecules of the present invention. Isolated nucleic acid molecules according to the present invention further include such molecules produced synthetically, purified from or provided in cells containing such nucleic acids, where the nucleic acid exists in other than a naturally occurring form, quantitatively or qualitatively.

"Isolated" used in reference to at least one polypeptide of the invention describes a state of isolation such that the peptide or polypeptide is not in a naturally occurring form and/or has been purified to remove at least some portion of cellular or non-cellular molecules with which the protein is naturally associated. However, "isolated" may include the addition of other functional or structural polypeptides for a specific purpose, where the other peptide may occur naturally associated with at least

one polypeptide of the present invention, but for which the resulting compound or composition does not exist naturally.

A "nucleic acid probe," "oligonucleotide probe," or "probe" as used herein comprises at least one detectably labeled or unlabeled nucleic acid which hybridizes under specified hybridization conditions with at least one other nucleic acid. This term also refers to a single- or partially double-stranded nucleic acid, oligonucleotide or polynucleotide that will associate with a complementary or partially complementary target nucleic acid to form at least a partially double-stranded nucleic acid molecule. A nucleic acid probe may be an oligonucleotide or a nucleotide polymer. A probe can optionally contain a detectable moiety which may be attached to the end(s) of the probe or be internal to the sequence of the probe, termed a "detectable probe" or "detectable nucleic acid probe."

The term "plasmid" refers to an extrachromosomal genetic element. The starting plasmids herein are either commercially available, publicly available on an unrestricted basis, or can be constructed from available plasmids in accordance with published procedures. In addition, equivalent plasmids to those described are known in the art and will be apparent to the ordinarily skilled artisan.

A "primer" is a nucleic acid fragment or oligonucleotide which functions as an initiating substrate for enzymatic or synthetic elongation of, for example, a nucleic acid molecule, e.g., using an amplification reaction, such as, but not limited to, a polymerase chain reaction (PCR), as known in the art.

"Recombinant DNA cloning vector" as used herein refers to any autonomously replicating agent, including, but not limited to, plasmids and phages, comprising a DNA molecule to which one or more additional DNA segments can or have been added.

The term "recombinant DNA expression vector" or "expression vector" as used herein refers to any recombinant DNA cloning vector, for example a plasmid or phage, in which a promoter and other regulatory elements are present thereby enabling transcription of an inserted DNA.

The term "stringency" refers to hybridization conditions for nucleic acids in solution. High stringency conditions disfavor non-homologous base pairing. Low stringency conditions have much less of this effect. Stringency may be altered, for example, by changes in temperature and/or salt concentration, or other conditions, as well known in the art.

A non-limiting example of "high stringency" conditions includes, for example, (a) a temperature of about 42°C , a formamide concentration of about ≤ 20%, and a low salt (SSC) concentration, or, alternatively, a temperature of about 65° C, or less, and a low salt (SSPE) concentration; (b) hybridization in 0.5 M NaHPO4, 7% sodium dodecyl sulfate (SDS), 1 mM EDTA at 65°C (See, e.g., Ausubel, *et al.*, ed., *Current Protocols in Molecular Biology*, 1987-1998, Wiley Interscience, New York, at §2.10.3). "SSC" comprises a hybridization and wash solution. A stock 20X SSC solution contains 3M sodium chloride, 0.3M sodium citrate, pH 7.0. "SSPE" comprises a hybridization and wash solution. A 1X

SSPE solution contains 180 mM NaCl, 9mM Na2HPO4, 0.9 mM NaH2PO4 and 1 mM EDTA, pH 7.4.

The term "transgene," as used herein, means a gene which is incorporated into the genome of an animal and is expressed in the animal, resulting in the presence of at least one huHDGFh polypeptide expressed by the transgenic animal.

The term "variant" when used in herein refers to an amino acid sequence that is altered by one or more amino acids. The variant may have "conservative" changes and/or "non-conservative" changes. Analagous minor variations may also include amino acid deletions or insertions, or both. Guidance in determining which amino acid residues may be substituted, inserted, or deleted without abolishing biological or immunological activity may be found using computer programs well known in the art, for example, DNASTAR software.

The term "vector" as used herein refers to a nucleic acid compound used for introducing exogenous or endogenous nucleic acid into host cells. A vector comprises a nucleotide sequence which may encode one or more polypeptide molecules. Plasmids, cosmids, viruses and bacteriophages, in a natural state or which have undergone recombinant engineering, are non-limiting examples of commonly used vectors to provide recombinant vectors comprising at least one desired isolated nucleic acid molecule.

Nucleic Acid Molecules

Using the information provided herein, such as the nucleotide sequences encoding at least 90-100% of the contiguous amino acids of SEQ ID NO:2, fragments or variants

thereof, or a vector comprising these sequences, a nucleic acid molecule of the present invention encoding a huHDGFh polypeptide can be obtained using well-known methods.

Nucleic acid molecules of the present invention can be in the form of RNA, such as mRNA, hnRNA, tRNA or any other form, or in the form of DNA, including, but not limited to, cDNA and genomic DNA obtained by cloning or produced synthetically, or any combination thereof. The DNA can be triple-stranded, double-stranded or single-stranded, or any combination thereof. Any portion of at least one strand of the DNA or RNA can be the coding strand, also known as the sense strand, or it can be the non-coding strand, also referred to as the anti-sense strand.

Isolated nucleic acid molecules of the present invention include nucleic acid molecules comprising an open reading frame (ORF) shown in SEQ ID NO:1, nucleic acid molecules comprising the coding sequence for a huHDGFh polypeptide; and nucleic acid molecules which comprise a nucleotide sequence substantially different from those described above but which, due to the degeneracy of the genetic code, still encode at least one huHDGFh polypeptide as described herein. Of course, the genetic code is well known in the art. Thus, it would be routine for one skilled in the art to generate such degenerate nucleic acid variants that code for specific huHDGFh polypeptides of the present invention. See, e.g., Ausubel, *et al.*, *supra*, and such nucleic acid variants are included in the present invention.

In a further embodiment, nucleic acid molecules are provided encoding the mature huHDGFh polypeptide or the full-length huHDGFh polypeptide lacking the N-terminal

methionine. The invention also provides an isolated nucleic acid molecule having the nucleotide sequence shown in SEQ ID NO:1, or a nucleic acid molecule having a sequence complementary thereto. Such isolated molecules, particularly nucleic acid molecules, are useful as probes for gene mapping by in situ hybridization with chromosomes, and for detecting transcription, translation and/or expression of the huHDGFh gene in human tissue, for instance, by Northern blot analysis for mRNA detection.

Unless otherwise indicated, all nucleotide sequences identified by sequencing a nucleic acid molecule herein can be or were identified using an automated nucleic acid sequencer, and all amino acid sequences of polypeptides encoded by nucleic acid molecules identified herein can be or were identified by codon correspondence or by translation of a nucleic acid sequence identified using method steps as described herein or as known in the art. Therefore, as is well known in the art that for any nucleic acid sequence identified by this automated approach, any nucleotide sequence identified herein may contain some errors which are reproducibly correctable by resequencing based upon an available or a vector or host cell containing the nucleic acid molecule using well-known methods.

Nucleotide sequences identified by automation are typically at least about 95% to at least about 99.999% identical to the actual nucleotide sequence of the sequenced nucleic acid molecule. The actual sequence can be more precisely identified by other approaches including manual nucleic acid sequencing methods well known in the art. As is also known in the art, a single insertion or deletion in

an identified nucleotide sequence compared to the actual sequence will cause a frame shift in translation of the nucleotide sequence such that the identified amino acid sequence encoded by an identified nucleotide sequence will be completely different from the amino acid sequence actually encoded by the sequenced nucleic acid molecule, beginning at the point of such an insertion or deletion.

Nucleic Acid Fragments

The present invention is further directed to fragments of the isolated nucleic acid molecules described herein. By a fragment of an isolated nucleic acid molecule is meant a molecule having at least 10 nucleotides of a nucleotide sequence of a cDNA or a nucleotide sequence shown in SEQ ID NO:1, and is intended to mean fragments at least about 10 nucleotides, and at least about 40 nucleotides in length, which are useful, *inter alia* as diagnostic probes and primers as described herein. Of course, larger fragments such as at least about 50, 100, 120, 200, 500, 1000, 1500, 2000, 2500, 3000, 3500, and/or 4000 or more nucleotides in length, are also useful according to the present invention as are fragments corresponding to most, if not all, of the nucleotide sequence as shown in SEQ ID NO:1. By a fragment at least 10 nucleotides in length, for example, is intended fragments which include 10 or more contiguous nucleotides from the nucleotide sequence of a cDNA or the nucleotide sequence as shown in SEQ ID NO:1, or consensus sequences thereof, as determined by methods known in the art.

Such nucleotide fragments are useful according to the present invention for screening DNA sequences that code for

one or more fragments of a huHDGFh polypeptide as described herein. Such screening, as a non-limiting example can include the use of so-called "DNA chips" for screening DNA sequences of the present invention of varying lengths, as described, e.g., in U.S. Patent Nos. 5,631,734, 5,624,711, 5,744,305, 5,770,456, 5,770,722, 5,675,443, 5,695,940, 5,710,000, 5,733,729, which are entirely incorporated herein by reference.

As indicated, nucleic acid molecules of the present invention which comprise a nucleic acid encoding a huHDGFh polypeptide can include, but are not limited to, those encoding the amino acid sequence of the mature polypeptide, by itself; the coding sequence for the mature polypeptide and additional sequences, such as the coding sequence of at least one signal leader or fusion peptide or of the mature polypeptide, with or without the aforementioned additional coding sequences, such as at least one intron, together with additional, non-coding sequences, including but not limited to, introns and non-coding 5' and 3' sequences, such as the transcribed, non-translated sequences that play a role in transcription, mRNA processing, including splicing and polyadenylation signals (for example - ribosome binding and stability of mRNA); an additional coding sequence which codes for additional amino acids, such as those which provide additional functionalities. Thus, the sequence encoding a polypeptide can be fused to a marker sequence, such as a sequence encoding a peptide that facilitates purification of the fused polypeptide.

Preferred nucleic acid fragments of the present invention also include nucleic acid molecules encoding epitope-bearing portions of a huHDGFh polypeptide.

Oligonucleotide and Polynucleotide Probes and/or Primers

In another aspect, the invention provides a polynucleotide (either DNA or RNA) that comprises at least about 20 nt, still more preferably at least about 30 nt, and even more preferably at least about 30-2000 nt of a nucleic acid molecule described herein. These are useful as diagnostic probes and primers as discussed above and in more detail below.

By a portion of a polynucleotide of "at least 10 nt in length," for example, is intended 10 or more contiguous nucleotides from the nucleotide sequence of the reference polynucleotide (e.g., at least one nucleic acid or at least one nucleotide sequence as shown in SEQ ID NO:1.

Of course, a polynucleotide which hybridizes only to a poly-A sequence (such as the 3' terminal poly(A) of a huHDGFh cDNA shown in SEQ ID NO:1, or to a complementary stretch of T (or U) resides, would not be included in a probe of the invention, since such a polynucleotide would hybridize to any nucleic acid molecule containing a poly (A) stretch or the complement thereof (e.g., practically any double-stranded cDNA clone).

The present invention also provides subsequences of full-length nucleic acids. Any number of subsequences can be obtained by reference to SEQ ID NO:1, or a complementary sequence, and using primers which selectively amplify, under stringent conditions to: at least two sites to the

polynucleotides of the present invention, or to two sites within the nucleic acid which flank and comprise a polynucleotide of the present invention, or to a site within a polynucleotide of the present invention and a site within the nucleic acid which comprises it. A variety of methods for obtaining 5′ and/or 3′ ends is well known in the art. See, e.g., RACE (Rapid Amplification of Complementary Ends) as described in M. A. Frohman, PCR Protocols: A Guide to Methods and Applications, M. A. Innis, D. H. Gelfand, J. J. Sninsky, T. J. White, Eds., Academic Press, Inc., San Diego, CA, pp. 28-38 (1990); see also, U.S. Patent No. 5,470,722, and Ausubel, *et al.*, *Current Protocols in Molecular Biology*, Unit 15.6, Eds., Greene Publishing and Wiley-Interscience, New York (1989-1998). Thus, the present invention provides huHDGFh polynucleotides having the sequence of the huHDGFh gene, nuclear transcript, cDNA, or complementary sequences and/or subsequences thereof.

Primer sequences can be obtained by reference to a contiguous subsequence of a polynucleotide of the present invention. Primers are chosen to selectively hybridize, under PCR amplification conditions, to a polynucleotide of the present invention in an amplification mixture comprising a genomic and/or cDNA library from the same species. Generally, the primers are complementary to a subsequence of the amplified nucleic acid. In some embodiments, the primers will be constructed to anneal at their 5′ terminal ends to the codon encoding the carboxy or amino terminal amino acid residue (or the complements thereof) of the polynucleotides of the present invention. The primer length in nucleotides is selected from the group of integers consisting of from at

least 15 to 50. Thus, the primers can be at least 15, 18, 20, 25, 30, 40, or 50 nucleotides in length or any range or value therein. A non-annealing sequence at the 5′ end of the primer (a "tail") can be added, for example, to introduce a cloning site at the terminal ends of the amplified DNA.

The amplification primers may optionally be elongated in the 3′ direction with additional contiguous or complementary nucleotides from the polynucleotide sequences, such as SEQ ID NO:1, from which they are derived. The number of nucleotides by which the primers can be elongated is selected from the group of integers consisting of from at least 1 to at least 25. Thus, for example, the primers can be elongated with an additional 1, 5, 10, or 15 nucleotides or any range or value therein. Those of skill will recognize that a lengthened primer sequence can be employed to increase specificity of binding (i.e., annealing) to a target sequence, or to add useful sequences, such as links or restriction sites.

The amplification products can be translated using expression systems well known to those of skill in the art and as discussed, infra. The resulting translation products can be confirmed as polypeptides of the present invention by, for example, assaying for the appropriate catalytic activity (e.g., specific activity and/or substrate specificity), or verifying the presence of one or more linear epitopes which are specific to a polypeptide of the present invention. Methods for protein synthesis from PCR derived templates are known in the art and available commercially. See, e.g., Amersham Life Sciences, Inc., Catalog '97, p. 354.

Polynucleotides Which Selectively Hybridize to a Polynucleotide as Described Herein

The present invention provides isolated nucleic acids that hybridize under selective hybridization conditions to a polynucleotide disclosed herein, e.g., SEQ ID NO:1. Thus, the polynucleotides of this embodiment can be used for isolating, detecting, and/or quantifying nucleic acids comprising such polynucleotides. For example, polynucleotides of the present invention can be used to identify, isolate, or amplify partial or full-length clones in a library. In some embodiments, the polynucleotides are genomic or cDNA sequences isolated, or otherwise complementary to, a cDNA from a human or mammalian nucleic acid library.

Preferably, the cDNA library comprises at least 80% full-length sequences, preferably at least 85% or 90% full-length sequences, and more preferably at least 95% full-length sequences. The cDNA libraries can be normalized to increase the representation of rare sequences. Low stringency hybridization conditions are typically, but not exclusively, employed with sequences having a reduced sequence identity relative to complementary sequences. Moderate and high stringency conditions can optionally be employed for sequences of greater identity. Low stringency conditions allow selective hybridization of sequences having about 70% sequence identity and can be employed to identify orthologous or paralogous sequences.

Optionally, polynucleotides of this invention will encode an epitope of a polypeptide encoded by the polynucleotides described herein. The polynucleotides of this

invention embrace nucleic acid sequences which can be employed for selective hybridization to a polynucleotide encoding a polypeptide of the present invention.

Screening polypeptides for specific binding to antibodies or fragments can be conveniently achieved using peptide display libraries. This method involves the screening of large collections of peptides for individual members having the desired function or structure. Antibody screening of peptide display libraries is well known in the art. The displayed peptide sequences can be from 3 to 5000 or more amino acids in length, frequently from 5-100 amino acids long, and often from about 8 to 15 amino acids long. In addition to direct chemical synthetic methods for generating peptide libraries, several recombinant DNA methods have been described. One type involves the display of a peptide sequence on the surface of a bacteriophage or cell. Each bacteriophage or cell contains the nucleotide sequence encoding the particular displayed peptide sequence. Such methods are described in PCT Patent Publication Nos. 91/17271, 91/18980, 91/19818, and 93/08278. Other systems for generating libraries of peptides have aspects of both in vitro chemical synthesis and recombinant methods. See, PCT Patent Publication Nos. 92/05258, 92/14843, and 96/19256. See also, U.S. Patent Nos. 5,658,754; and 5,643,768. Peptide display libraries, vector, and screening kits are commercially available from such suppliers as Invitrogen (Carlsbad, CA). (See, e.g., Ausubel, *supra*; or Sambrook, *supra*).

Polynucleotides Complementary to the Polynucleotides

As indicated above, the present invention provides isolated nucleic acids comprising huHDGFh polynucleotides, wherein the polynucleotides are complementary to the polynucleotides described herein, above. As those of skill in the art will recognize, complementary sequences base-pair throughout the entirety of their length with such polynucleotides (i.e., have 100% sequence identity over their entire length). Complementary bases associate through hydrogen bonding in double-stranded nucleic acids. For example, the following base pairs are complementary: guanine and cytosine; adenine and thymine; and adenine and uracil. (See, e.g., Ausubel, *supra*; or Sambrook, *supra*)

Construction of Nucleic Acids

The isolated nucleic acids of the present invention can be made using (a) standard recombinant methods, (b) synthetic techniques, (c) purification techniques, or combinations thereof, as well known in the art.

The nucleic acids may conveniently comprise sequences in addition to a polynucleotide of the present invention. For example, a multi-cloning site comprising one or more endonuclease restriction sites may be inserted into the nucleic acid to aid in isolation of the polynucleotide. Also, translatable sequences may be inserted to aid in the isolation of the translated polynucleotide of the present invention. For example, a hexa-histidine marker sequence provides a convenient means to purify the proteins of the present invention. The nucleic acid of the present invention - excluding the polynucleotide sequence - is optionally a

vector, adapter, or linker for cloning and/or expression of a polynucleotide of the present invention.

Additional sequences may be added to such cloning and/or expression sequences to optimize their function in cloning and/or expression, to aid in isolation of the polynucleotide, or to improve the introduction of the polynucleotide into a cell. Typically, the length of a nucleic acid of the present invention less the length of its polynucleotide of the present invention is less than 20 kilobase pairs, often less than 15 kb, and frequently less than 10 kb. Use of cloning vectors, expression vectors, adapters, and linkers is well known in the art. (See, e.g., Ausubel, *supra*; or Sambrook, *supra*)

Recombinant Methods for Constructing Nucleic Acids

The isolated nucleic acid compositions of this invention, such as RNA, cDNA, genomic DNA, or a hybrid thereof, can be obtained from biological sources using any number of cloning methodologies known to those of skill in the art. In some embodiments, oligonucleotide probes which selectively hybridize, under stringent conditions, to the polynucleotides of the present invention are used to identify the desired sequence in a cDNA or genomic DNA library. While isolation of RNA, and construction of cDNA and genomic libraries is well known to those of ordinary skill in the art. (See, e.g., Ausubel, *supra*; or Sambrook, *supra*)

Nucleic Acid Screening and Isolation Methods

A cDNA or genomic library can be screened using a probe based upon the sequence of a polynucleotide of the present

invention, such as those disclosed herein. Probes may be used to hybridize with genomic DNA or cDNA sequences to isolate homologous genes in the same or different organisms. Those of skill in the art will appreciate that various degrees of stringency of hybridization can be employed in the assay; and either the hybridization or the wash medium can be stringent. As the conditions for hybridization become more stringent, there must be a greater degree of complementarity between the probe and the target for duplex formation to occur. The degree of stringency can be controlled by temperature, ionic strength, pH and the presence of a partially denaturing solvent such as formamide. For example, the stringency of hybridization is conveniently varied by changing the polarity of the reactant solution through, for example, manipulation of the concentration of formamide within the range of 0% to 50%. The degree of complementarity (sequence identity) required for detectable binding will vary in accordance with the stringency of the hybridization medium and/or wash medium. The degree of complementarity will optimally be 100%; however, it should be understood that minor sequence variations in the probes and primers may be compensated for by reducing the stringency of the hybridization and/or wash medium.

Methods of amplification of RNA or DNA are well known in the art and can be used according to the present invention without undue experimentation, based on the teaching and guidance presented herein.

Known methods of DNA or RNA amplification include, but are not limited to, polymerase chain reaction (PCR) and related amplification processes (see, e.g., U.S. Patent Nos.

4,683,195, 4,683,202, 4,800,159, 4,965,188, to Mullis, *et al.*; 4,795,699 and 4,921,794 to Tabor, *et al*; 5,142,033 to Innis; 5,122,464 to Wilson, *et al.*; 5,091,310 to Innis; 5,066,584 to Gyllensten, *et al*; 4,889,818 to Gelfand, *et al*; 4,994,370 to Silver, *et al*; 4,766,067 to Biswas; 4,656,134 to Ringold) and RNA mediated amplification which uses anti-sense RNA to the target sequence as a template for double-stranded DNA synthesis (U.S. Patent No. 5,130,238 to Malek, *et al*, with the tradename NASBA), the entire contents of which are herein incorporated by reference. (See, e.g., Ausubel, *supra*; or Sambrook, *supra*)

For instance, polymerase chain reaction (PCR) technology can be used to amplify the sequences of polynucleotides of the present invention and related genes directly from genomic DNA or cDNA libraries. PCR and other in vitro amplification methods may also be useful, for example, to clone nucleic acid sequences that code for proteins to be expressed, to make nucleic acids to use as probes for detecting the presence of the desired mRNA in samples, for nucleic acid sequencing, or for other purposes. Examples of techniques sufficient to direct persons of skill through in vitro amplification methods are found in Berger, Sambrook, and Ausubel, as well as Mullis, *et al.*, U.S. Patent No. 4,683,202 (1987); and Innis, *et al.*, PCR Protocols A Guide to Methods and Applications, Eds., Academic Press Inc., San Diego, CA (1990). Commercially available kits for genomic PCR amplification are known in the art. See, e.g., Advantage-GC Genomic PCR Kit (Clontech). The T4 gene 32 protein (Boehringer Mannheim) can be used to improve yield of long PCR products.

Synthetic Methods for Constructing Nucleic Acids

The isolated nucleic acids of the present invention can also be prepared by direct chemical synthesis by methods such as the phosphotriester method of Narang, *et al.*, *Meth. Enzymol.* 68:90-99 (1979); the phosphodiester method of Brown, *et al.*, *Meth. Enzymol.* 68:109-151 (1979); the diethylphosphoramidite method of Beaucage, *et al.*, *Tetra. Letts.* 22:1859-1862 (1981); the solid phase phosphoramidite triester method described by Beaucage and Caruthers, Tetra. Letts. 22(20):1859-1862 (1981), e.g., using an automated synthesizer, e.g., as described in Needham-VanDevanter, *et al.*, *Nucleic Acids Res.* 12:6159-6168 (1984); and the solid support method of U.S. Patent No. 4,458,066. Chemical synthesis generally produces a single-stranded oligonucleotide, which may be converted into double-stranded DNA by hybridization with a complementary sequence, or by polymerization with a DNA polymerase using the single strand as a template. One of skill in the art will recognize that while chemical synthesis of DNA can be limited to sequences of about 100 or more bases, longer sequences may be obtained by the ligation of shorter sequences.

Recombinant Expression Cassettes

The present invention further provides recombinant expression cassettes comprising a nucleic acid of the present invention. A nucleic acid sequence of the present invention, for example a cDNA or a genomic sequence encoding a full-length polypeptide of the present invention, can be used to construct a recombinant expression cassette which can be

introduced into at least one desired host cell. A recombinant expression cassette will typically comprise a polynucleotide of the present invention operably linked to transcriptional initiation regulatory sequences that will direct the transcription of the polynucleotide in the intended host cell.

Both heterologous and non-heterologous (i.e., endogenous) promoters can be employed to direct expression of the nucleic acids of the present invention. These promoters can also be used, for example, in recombinant expression cassettes to drive expression of antisense nucleic acids to reduce, increase, or alter huHDGFh content and/or composition in a desired tissue.

In some embodiments, isolated nucleic acids which serve as promoter or enhancer elements can be introduced in the appropriate position (generally upstream) of a non-heterologous form of a polynucleotide of the present invention so as to up or down regulate expression of a polynucleotide of the present invention. For example, endogenous promoters can be altered *in vivo* or *in vitro* by mutation, deletion and/or substitution.

A polynucleotide of the present invention can be expressed in either sense or anti-sense orientation as desired. It will be appreciated that control of gene expression in either sense or anti-sense orientation can have a direct impact on the observable characteristics.

Another method of suppression is sense suppression. Introduction of nucleic acid configured in the sense orientation has been shown to be an effective means by which to block the transcription of target genes.

A variety of cross-linking agents, alkylating agents and radical generating species as pendant groups on polynucleotides of the present invention can be used to bind, label, detect and/or cleave nucleic acids. Knorre, *et al.*, *Biochimie* 67:785-789 (1985); Vlassov, *et al.*, *Nucleic Acids Res.* 14:4065-4076 (1986); Iverson and Dervan, *J. Am. Chem. Soc.* 109:1241-1243 (1987); Meyer, *et al.*, *J. Am. Chem. Soc.* 111:8517-8519 (1989); Lee, *et al.*, *Biochemistry* 27:3197-3203 (1988); Home, *et al.*, *J. Am. Chem. Soc.* 112:2435-2437 (1990); Webb and Matteucci, *J. Am. Chem. Soc.* 108:2764-2765 (1986); *Nucleic Acids Res.* 14:7661-7674 (1986); Feteritz, *et al.*, *J. Am. Chem. Soc.* 113:4000 (1991). Various compounds to bind, detect, label, and/or cleave nucleic acids are known in the art. See, for example, U.S. Patent Nos. 5,543,507; 5,672,593; 5,484,908; 5,256,648; and 5,681941, each entirely incorporated herein by reference.

Vectors and Host Cells

The present invention also relates to vectors that include isolated nucleic acid molecules of the present invention, host cells that are genetically engineered with the recombinant vectors, and the production of huHDGFh polypeptides or fragments thereof by recombinant techniques, as is well known in the art. See, e.g., Sambrook, *et al.*, 1989; Ausubel, *et al.*, 1987-1998, each entirely incorporated herein by reference.

The polynucleotides can optionally be joined to a vector containing a selectable marker for propagation in a host. Generally, a plasmid vector is introduced in a precipitate, such as a calcium phosphate precipitate, or in

a complex with a charged lipid. If the vector is a virus, it can be packaged in vitro using an appropriate packaging cell line and then transduced into host cells.

The DNA insert should be operatively linked to an appropriate promoter, such as the phage lambda PL promoter, the *E. coli lac, trp* and *tac* promoters, the SV40 early and late promoters and promoters of retroviral LTRs, or any other suitable promoter. Other suitable promoters will be known to the skilled artisan. The expression constructs will further contain sites for transcription initiation, termination and, in the transcribed region, a ribosome binding site for translation. The coding portion of the mature transcripts expressed by the constructs will preferably include a translation initiating at the beginning and a termination codon (e.g., UAA, UGA or UAG) appropriately positioned at the end of the mRNA to be translated, with VAA and VAG preferred for mammalian or eukaryotic cell expression.

Expression vectors will preferably include at least one selectable marker. Such markers include, e.g., dihydrofolate reductase, ampicillin (G418), or neomycin resistance for eukaryotic cell culture, and tetracycline or ampicillin resistance genes for culturing in E. coli and other bacteria or prokaryotics. Representative examples of appropriate hosts include, but are not limited to, bacterial cells, such as E. coli, Streptomyces and Salmonella typhimurium cells; fungal cells, such as yeast cells; insect cells such as Drosophila S2 and Spodoptera Sf9 cells; animal cells such as CHO, COS and Bowes melanoma cells; and plant cells. Appropriate culture mediums and conditions for the

above-described host cells are known in the art. Vectors preferred for use in bacteria include pQE70, pQE60 and pQE-9, available from Qiagen; pBS vectors, Phagescript vectors, Bluescript vectors, pNH8A, pNH16a, pNH18A, pNH46A, available from Stratagene; and ptrc99a, pKK223-3, pKK233-3, pDR540, pRIT5 available from Pharmacia. Preferred eucaryotic vectors include pWLNEO, pSV2CAT, pOG44, pXT1 and pSG available from Stratagene; and pSVK3, pBPV, pMSG and pSVL available from Pharmacia. Other suitable vectors will be readily apparent to the skilled artisan.

Introduction of a vector construct into a host cell can be effected by calcium phosphate transfection, DEAE-dextran mediated transfection, cationic lipid-mediated transfection, electroporation, transduction, infection or other methods. Such methods are described in many standard laboratory manuals, such as Sambrook, supra, Chapters 1-4 and 16-18; Ausubel, supra, Chapters 1, 9, 13, 15, 16.

Polypeptide(s) of the present invention can be expressed in a modified form, such as a fusion protein, and can include not only secretion signals, but also additional heterologous functional regions. For instance, a region of additional amino acids, particularly charged amino acids, can be added to the N-terminus of a polypeptide to improve stability and persistence in the host cell, during purification, or during subsequent handling and storage. Also, peptide moieties can be added to a polypeptide to facilitate purification. Such regions can be removed prior to final preparation of a polypeptide. Such methods are described in many standard laboratory manuals, such as

Sambrook, supra, Chapters 17.29-17.42 and 18.1-18.74; Ausubel, supra, Chapters 16, 17 and 18.

Expression of Proteins in Host Cells

Using nucleic acids of the present invention, one may express a protein of the present invention in a recombinantly engineered cell, such as bacteria, yeast, insect, or mammalian cells. The cells produce the protein in a non-natural condition (e.g., in quantity, composition, location, and/or time), because they have been genetically altered through human intervention to do so.

It is expected that those of skill in the art are knowledgeable in the numerous expression systems available for expression of a nucleic acid encoding a protein of the present invention. No attempt to describe in detail the various methods known for the expression of proteins in prokaryotes or eukaryotes will be made.

In brief summary, the expression of isolated nucleic acids encoding a protein of the present invention will typically be achieved by operably linking, for example, the DNA or cDNA to a promoter (which is either constitutive or inducible), followed by incorporation into an expression vector. The vectors can be suitable for replication and integration in either prokaryotes or eukaryotes. Typical expression vectors contain transcription and translation terminators, initiation sequences and promoters useful for regulation of the expression of the DNA encoding a protein of the present invention. To obtain high level expression of a cloned gene, it is desirable to construct expression vectors which contain, at the minimum, a strong promoter to direct

transcription, a ribosome binding site for translational initiation, and a transcription/translation terminator. One of skill would recognize that modifications can be made to a protein of the present invention without diminishing its biological activity. Some modifications may be made to facilitate the cloning, expression, or incorporation of the targeting molecule into a fusion protein. Such modifications are well known to those of skill in the art and include, for example, a methionine added at the amino terminus to provide an initiation site, or additional amino acids (e.g., poly His) placed on either terminus to create conveniently located restriction sites or termination codons or purification sequences.

Alternatively, nucleic acids of the present invention can be expressed in a host cell by turning on (by manipulation) in a host cell that contains endogenous DNA encoding a polypeptide of the present invention. Such methods are well known in the art, e.g., as described in US patent Nos. 5,580,734, 5,641,670, 5,733,746, and 5,733,761, entirely incorporated herein by reference.

Expression in Prokaryotes

Prokaryotic cells may be used as hosts for expression. Prokaryotes most frequently are represented by various strains of E. coli; however, other microbial strains may also be used. Commonly used prokaryotic control sequences which are defined herein to include promoters for transcription initiation, optionally with an operator, along with ribosome binding site sequences, include such commonly used promoters as the beta lactamase (penicillinase) and lactose (lac)

promoter systems (Chang, *et al.*, *Nature* 198:1056 (1977)), the tryptophan (trp) promoter system (Goeddel, *et al.*, *Nucleic Acids Res.* 8:4057 (1980)) and the lambda derived P L promoter and N-gene ribosome binding site (Shimatake, *et al.*, *Nature* 292:128 (1981)). The inclusion of selection markers in DNA vectors transfected in E. coli is also useful. Examples of such markers include genes specifying resistance to ampicillin, tetracycline, or chloramphenicol.

The vector is selected to allow introduction into the appropriate host cell. Bacterial vectors are typically of plasmid or phage origin. Appropriate bacterial cells are infected with phage vector particles or transfected with naked phage vector DNA. If a plasmid vector is used, the bacterial cells are transformed with the plasmid vector DNA. Expression systems for expressing a protein of the present invention are available using Bacillus sp. and Salmonella (Palva, *et al.*, *Gene* 22:229-235 (1983); Mosbach, *et al.*, *Nature* 302:543-545 (1983)).

Expression in Eukaryotes

A variety of eukaryotic expression systems such as yeast, insect cell lines, plant and mammalian cells, are known to those of skill in the art. As explained briefly below, a nucleic acid of the present invention can be expressed in these eukaryotic systems.

Synthesis of heterologous proteins in yeast is well known. F. Sherman, *et al.*, *Methods in Yeast Genetics*, Cold Spring Harbor Laboratory (1982) is a well-recognized work describing the various methods available to produce the protein in yeast. Two widely utilized yeast for production

of eukaryotic proteins are Saccharomyces cerevisiae and Pichia pastoris. Vectors, strains, and protocols for expression in Saccharomyces and Pichia are known in the art and available from commercial suppliers (e.g., Invitrogen). Suitable vectors usually have expression control sequences, such as promoters, including 3-phosphoglycerate kinase or alcohol oxidase, and an origin of replication, termination sequences and the like as desired.

A protein of the present invention, once expressed, can be isolated from yeast by lysing the cells and applying standard protein isolation techniques to the lysates. The monitoring of the purification process can be accomplished by using Western blot techniques or radioimmunoassay of other standard immunoassay techniques.

The sequences encoding proteins of the present invention can also be ligated to various expression vectors for use in transfecting cell cultures of, for instance, mammalian, insect, or plant origin. Illustrative of cell cultures useful for the production of the peptides are mammalian cells. Mammalian cell systems often will be in the form of monolayers of cells although mammalian cell suspensions may also be used. A number of suitable host cell lines capable of expressing intact proteins have been developed in the art, and include the HEK293, BHK21, and CHO cell lines. Expression vectors for these cells can include expression control sequences, such as an origin of replication, a promoter (e.g., the CMV promoter, a HSV tk promoter or pgk (phosphoglycerate kinase) promoter), an enhancer (Queen, *et al.*, *Immunol. Rev.* 89:49 (1986)), and processing information sites, such as ribosome binding sites, RNA splice sites,

polyadenylation sites (e.g., an SV40 large T Ag poly A addition site), and transcriptional terminator sequences. Other animal cells useful for production of proteins of the present invention are available, for instance, from the American Type Culture Collection Catalogue of Cell Lines and Hybridomas (7th edition, 1992).

Appropriate vectors for expressing proteins of the present invention in insect cells are usually derived from the SF9 baculovirus. Suitable insect cell lines include mosquito larvae, silkworm, armyworm, moth and Drosophila cell lines such as a Schneider cell line (See Schneider, J. *Embryol. Exp. Morphol.* 27:353-365 (1987).

As with yeast, when higher animal or plant host cells are employed, polyadenlyation or transcription terminator sequences are typically incorporated into the vector. An example of a terminator sequence is the polyadenlyation sequence from the bovine growth hormone gene. Sequences for accurate splicing of the transcript may also be included. An example of a splicing sequence is the VP1 intron from SV40 (Sprague, *et al., J. Virol.* 45:773-781 (1983)). Additionally, gene sequences to control replication in the host cell may be incorporated into the vector such as those found in bovine papilloma virus type-vectors. M. Saveria-Campo, Bovine Papilloma Virus DNA, a Eukaryotic Cloning Vector in DNA Cloning Vol. II, a Practical Approach, D. M. Glover, Ed., IRL Press, Arlington, VA, pp. 213-238 (1985).

Protein Purification

A huHDGFh polypeptide can be recovered and purified from recombinant cell cultures by well-known methods

including ammonium sulfate or ethanol precipitation, acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, hydroxylapatite chromatography and lectin chromatography. Most preferably, high performance liquid chromatography ("HPLC") is employed for purification. Polypeptides of the present invention include naturally purified products, products of chemical synthetic procedures, and products produced by recombinant techniques from a prokaryotic or eucaryotic host, including, for example, bacterial, yeast, higher plant, insect and mammalian cells. Depending upon the host employed in a recombinant production procedure, the polypeptides of the present invention can be glycosylated or can be non-glycosylated. In addition, polypeptides of the invention can also include an initial modified methionine residue, in some cases as a result of host-mediated processes. Such methods are described in many standard laboratory manuals, such as Sambrook, supra, Chapters 17.37-17.42; Ausubel, supra, Chapters 10, 12, 13, 16, 18 and 20.

huHDGFh POLYPEPTIDES AND FRAGMENTS AND VARIANTS

The invention further provides an isolated huHDGFh polypeptide having fragments or specified variants of the amino acid sequence encoded by the amino acid sequence in SEQ ID NO:2.

The isolated proteins of the present invention comprise a polypeptide encoded by any one of the polynucleotides of the present invention as discussed more fully, supra, or

polypeptides which are specified fragments or variants thereof.

An exemplary polypeptide sequence is provided in SEQ ID NO:2. The proteins of the present invention or variants thereof can comprise any number of contiguous amino acid residues from a polypeptide of the present invention, wherein that number is selected from the group of integers consisting of from 90-100% of the number of contiguous residues in a full-length huHDGFh polypeptide. Optionally, this subsequence of contiguous amino acids is at least 50, 60, 70, 80, or 90 amino acids in length. Further, the number of such subsequences can be any integer selected from the group consisting of from 1 to 20, such as 2, 3, 4, or 5.

As those of skill will appreciate, the present invention includes biologically active polypeptides of the present invention (i.e., enzymes). Biologically active polypeptides have a specific activity at least 20%, 30%, or 40%, and preferably at least 50%, 60%, or 70%, and most preferably at least 80%, 90%, or 95%-100% of that of the native (non-synthetic) or endogenous polypeptide.

Generally, the polypeptides of the present invention will, when presented as an immunogen, elicit production of an antibody specifically reactive to a polypeptide of the present invention encoded by a polynucleotide of the present invention as described, supra. Exemplary polypeptides include those which are full-length, such as those disclosed herein. Further, the proteins of the present invention will not bind to antisera raised against a polypeptide of the present invention which has been fully immunosorbed with the same polypeptide. Immunoassays for determining binding are

well known to those of skill in the art. A preferred immunoassay is a competitive immunoassay as discussed, infra. Thus, the proteins of the present invention can be employed as immunogens for constructing antibodies immunoreactive to a protein of the present invention for such exemplary utilities as immunoassays or protein purification techniques.

A huHDGFh polypeptide of the present invention can include one or more amino acid substitutions, deletions or additions, either from natural mutations or human manipulation, as specified herein.

Of course, the number of amino acid substitutions a skilled artisan would make depends on many factors, including those described above. Generally speaking, the number of amino acid substitutions, insertions or deletions for any given huHDGFh polypeptide will not be more than 40, 30, 20, 10, 5, or 3, such as 1-30 or any range or value therein, as specified herein.

Amino acids in a huHDGFh polypeptide of the present invention that are essential for function can be identified by methods known in the art, such as site-directed mutagenesis or alanine-scanning mutagenesis (Cunningham and Wells, *Science* 244:1081-1085 (1989)). The latter procedure introduces single alanine mutations at every residue in the molecule. The resulting mutant molecules are then tested for biological activity. Sites that are critical for ligand-protein binding can also be identified by structural analysis such as crystallization, nuclear magnetic resonance or photoaffinity labeling (Smith, *et al.*, *J. Mol. Biol.* 224:899-904 (1992) and de Vos, *et al.*, *Science* 255:306-312 (1992)).

huHDGFh polypeptides of the present invention can include but are not limited to, at least one selected from SEQ ID NO:2.

A huHDGFh polypeptide can further comprise a polypeptide of 203 contiguous amino acids of SEQ ID NO:2 .

A huHDGFh polypeptide further includes an amino acid sequence selected from SEQ ID NO:2.

Non-limiting mutants that can enhance or maintain at least one of the listed activities include, but are not limited to, any of the above polypeptides, further comprising at least one mutation corresponding to at least one substitution selected from the group consisting of 9F, 11P, 15I, 29V, 31D, 32V, 32I, 33A, 34D, 41T, 41P, 44L, 62I, 66S, 66D, 67K, 68N, 68C, 70E, 72Y, 75P, 89D, 89Q, 93K, 93H, 94A, 95S, 96Y, 97S, 98S, 98A, 99Q, 99Y, 100P, 101P, 103T, 103S, 104K, 104S, and 105S of SEQ ID NO:2.

Antigenic/Epitope Comprising huHDGFh Peptide and Polypeptides

In another aspect, the invention provides a peptide or polypeptide comprising an epitope-bearing portion of a polypeptide of the invention according to methods well known in the art. See, e.g., Colligan, *et al*,. ed., *Current Protocols in Immunology*, Greene Publishing, NY (1993-1998), Ausubel, *supra*, each entirely incorporated herein by reference.

The epitope of this polypeptide portion is an immunogenic or antigenic epitope of a polypeptide described herein. An "immunogenic epitope" can be defined as a part of a polypeptide that elicits an antibody response when the

whole polypeptide is the immunogen. On the other hand, a region of a polypeptide molecule to which an antibody can bind is defined as an "antigenic epitope." The number of immunogenic epitopes of a polypeptide generally is less than the number of antigenic epitopes. See, for instance, Geysen, *et al.*, *Proc. Natl. Acad. Sci.* USA 81:3998-4002 (1983).

As to the selection of peptides or polypeptides bearing an antigenic epitope (i.e., that contain at least a portion of a region of a polypeptide molecule to which an antibody can bind), it is well known in the art that relatively short synthetic peptides that mimic part of a polypeptide sequence are routinely capable of eliciting an antiserum that reacts with the partially mimicked polypeptide. See, for instance, J. G. Sutcliffe, *et al.*, "Antibodies that react with pre-identified sites on polypeptides," *Science* 219:660-666 (1983).

Antigenic epitope-bearing peptides and polypeptides of the invention are useful to raise antibodies, including monoclonal antibodies, or screen antibodies, including fragments or single chain antibodies, that bind specifically to a polypeptide of the invention. See, for instance, Wilson, *et al.*, *Cell* 37:767-778 (1984) at 777. Antigenic epitope-bearing peptides and polypeptides of the invention preferably contain a sequence of at least five, more preferably at least nine, and most preferably between at least about 15 to about 30 amino acids contained within the amino acid sequence of a polypeptide of the invention.

The epitope-bearing peptides and polypeptides of the invention can be produced by any conventional means. R. A.

Houghten, "General method for the rapid solid-phase synthesis of large numbers of peptides: specificity of antigen-antibody interaction at the level of individual amino acids," *Proc. Natl. Acad. Sci.* USA 82:5131-5135 (1985). This "Simultaneous Multiple Peptide Synthesis (SMPS)" process is further described in U.S. Patent No. 4,631,211 to Houghten, *et al.* (1986).

As one of skill in the art will appreciate, huHDGFh polypeptides of the present invention and the epitope-bearing fragments thereof described above can be combined with parts of the constant domain of immunoglobulins (IgG), resulting in chimeric polypeptides. These fusion proteins facilitate purification and show an increased half-life in vivo. This has been shown, e.g., for chimeric proteins consisting of the first two domains of the human CD4-polypeptide and various domains of the constant regions of the heavy or light chains of mammalian immunoglobulins (EPA 394,827; Traunecker, *et al.*, *Nature* 331:84-86 (1988)). Fusion proteins that have a disulfide-linked dimeric structure due to the IgG part can also be more efficient in binding and neutralizing other molecules than the monomeric huHDGFh polypeptide or polypeptide fragment alone (Fountoulakis, *et al.*, *J. Biochem.* 270:3958-3964 (1995)).

Production of Antibodies

The polypeptides of this invention and fragments thereof may be used in the production of antibodies. The term "antibody" as used herein describes antibodies, fragments of antibodies (such as, but not limited, to Fab, Fab', Fab2', and Fv fragments), and modified versions

thereof, as well known in the art (e.g., chimeric, humanized, recombinant, veneered, resurfaced or CDR-grafted) such antibodies are capable of binding antigens of a similar nature as the parent antibody molecule from which they are derived. The instant invention also encompasses single chain polypeptide binding molecules.

The production of antibodies, both monoclonal and polyclonal, in animals is well known in the art. See, e.g., Colligan, supra, entirely incorporated herein by reference.

Single chain antibodies and libraries thereof are yet another variety of genetically engineered antibody technology that is well known in the art. (See, e.g., R. E. Bird, *et al.*, *Science* 242:423-426 (1988); PCT Publication Nos. WO 88/01649, WO 90/14430, and WO 91/10737. Single chain antibody technology involves covalently joining the binding regions of heavy and light chains to generate a single polypeptide chain. The binding specificity of the intact antibody molecule is thereby reproduced on a single polypeptide chain.

Antibodies included in this invention are useful in diagnostics, therapeutics or in diagnostic/therapeutic combinations.

The polypeptides of this invention or suitable fragments thereof can be used to generate polyclonal or monoclonal antibodies, and various inter-species hybrids, or humanized antibodies, or antibody fragments, or single-chain antibodies. The techniques for producing antibodies are well known to skilled artisans. (See, e.g., Colligan supra; Monoclonal Antibodies: Principles & Applications, Ed. J. R. Birch & E. S. Lennox, Wiley-Liss (1995).

A polypeptide used as an immunogen may be modified or administered in an adjuvant, by subcutaneous or intraperitoneal injection into, for example, a mouse or a rabbit. For the production of monoclonal antibodies, spleen cells from immunized animals are removed, fused with myeloma or other suitable known cells, and allowed to become monoclonal antibody producing hybridoma cells in the manner known to the skilled artisan. Hybridomas that secrete a desired antibody molecule can be screened by a variety of well known methods, for example ELISA assay, Western blot analysis, or radioimmunoassay (Lutz, *et al. Exp. Cell Res.* 175:109-124 (1988); Monoclonal Antibodies: Principles & Applications, Ed. J. R. Birch & E. S. Lennox, Wiley-Liss (1995); Colligan, supra).

For some applications labeled antibodies are desirable. Procedures for labeling antibody molecules are widely known, including for example, the use of radioisotopes, affinity labels, such as biotin or avidin, enzymatic labels, for example horseradish peroxidase, and fluorescent labels, such as FITC or rhodamine (See, e.g., Colligan, supra).

Labeled antibodies are useful for a variety of diagnostic applications. In one embodiment the present invention relates to the use of labeled antibodies to detect the presence of a huHDGFh polypeptide. Alternatively, the antibodies could be used in a screen to identify potential modulators of a huHDGFh polypeptide. For example, in a competitive displacement assay, the antibody or compound to be tested is labeled by any suitable method. Competitive displacement of an antibody from an antibody-antigen complex by a test compound such that a test compound-antigen complex

is formed provides a method for identifying compounds that bind HPLFP.

Transgenics and Chimeric Non-Human Mammals

Another embodiment of the present invention provides transgenic non-human mammals carrying a recombinant human huHDGFh gene construct in its somatic and germ cells. The recombinant gene construct may be composed of regulatory DNA sequences that belong to the native huHDGFh gene or those which are derived from an alternative source. These regulatory sequences are functionally linked to the human huHDGFh coding region, resulting in the constitutive and/or regulatable expression of human huHDGFh in the body of the transgenic non-human mammal. The most important of such regulatory sequences is the promoter. Promoters are defined in this context as any and all DNA elements necessary for the functional expression of a gene. Promoters drive the expression of structural genes and may be modulated by inducers and repressors. Numerous promoters have been described in the literature and are easily within the grasp of the ordinarily skilled artisan. Viral promoters, such as the SV40 early promoter, are consistent with the invention though mammalian promoters are preferred. The promoter is chosen such that the level of expression is sufficient to promote physiological consequences in the transgenic non-human mammal, or ancestor of said mammal. Preferably, the genome of the transgenic mammal contains at least 30 copies of a transgene. More preferably, the genome of the transgenic mammal contains at least 50 copies, and may contain 100-200 or more copies of the transgene.

Generally, said nucleic acid is introduced into said mammal at an embryonic stage, preferably the 1-1000 cell or oocyte stage, and, most preferably not later than about the 64-cell stage. Most preferably the transgenic mammal is homozygous for the transgene.

The techniques described in Leder, U.S. Patent No. 4,736,866 (hereby entirely incorporated by reference) for producing transgenic non-human mammals may be used for the production of a transgenic non-human mammal of the present invention. The various techniques described in U.S. patent Nos. 5,454,807, 5,073,490, 5,347,075 and 4,736,866, the entire contents of which are hereby incorporated by reference, may also be used. Such methods are also described in Methods in Molecular Biology, Vol. 18, 1993, Transgenesis Techniques, Principles and Protocols, (Murphy, D., and Carter, D.A.) as well as in U.S. Patents #5,174,986, #5,175,383, #5,175,384, and #5,175,385, all of which are herein incorporated by reference.

Also intended to be within the scope of the present invention are chimeric non-human mammals in which fewer than all of the somatic and germ cells contain a DNA construct comprising a nucleic acid encoding a huHDGFh polypeptide of the present invention. Contemplated chimeric non-human mammals include animals produced when fewer than all of the cells of the morula are transfected in the process of producing the transgenic animal.

Transgenic and chimeric non-human mammals having human cells or tissue engrafted therein are also encompassed by the present invention. Methods for providing chimeric non-human mammals are provided, e.g., in U.S. Serial Nos. 07/508,225,

07/518,748, 07/529,217, 07/562,746, 07/596,518, 07/574,748, 07/575,962, 07/207,273, 07/241,590 and 07/137,173, which are entirely incorporated herein by reference, for their description of how to engraft human cells or tissue into non-human mammals.

Alternatively, genetic constructs comprising at least one of the huHDGFh nucleic acid sequences as defined herein may be used to create transgenic "knockouts" of the huHDGFh gene. Accordingly, the present invention also provides a transgenic animal which has been engineered by homologous recombination to be deficient in the expression of the endogenous huHDGFh gene. Further, the invention provides a method of producing an heterozygous or homozygous transgenic animal deficient in or lacking functional HDGFH proteins, respectfully, said method comprising:

a) obtaining a DNA construct comprising a disrupted huHDGFh gene, wherein said disruption is by the insertion of an heterologous marker sequence;

b) introducing said DNA construct into an ES cell of said animal such that the endogenous HDGFH gene is disrupted by homologous recombination;

c) selecting ES cells comprising said disrupted allele;

d) incorporating the ES cells of step c) into a mouse embryo;

e) transferring said embryo into a pseudopregnant animal of the said species;

f) developing said embryo into a viable offspring;

g) screening offspring to identify heterozygous animal comprising said disrupted HDGFH gene; and

h) if desired, breeding said heterozygous animal to produce homozygous transgenic animals of said species, wherein the said homozygous animal does not express functional HDGFH proteins.

Transgenic and chimeric non-human mammals of the present invention may be used for analyzing the consequences of over-expression of at least one huHDGFh polypeptide *in vivo*. Such animals are also useful for testing the effectiveness of therapeutic and/or diagnostic agents, either associated or unassociated with delivery vectors or vehicles, which preferentially bind to an huHDGFh polypeptide of the present invention or act to indirectly modulate huHDGFh activity.

huHDGFh transgenic non-human mammals are useful as an animal models in both basic research and drug development endeavors. Transgenic animals carrying at least one huHDGFh polypeptide or nucleic acid can be used to test compounds or other treatment modalities which may prevent, suppress, or cure a pathology or disease associated with at least one of the above mentioned huHDGFh activities. Such transgenic animals can also serve as a model for the testing of diagnostic methods for those same diseases. Furthermore, tissues derived from huHDGFh transgenic non-human mammals are useful as a source of cells for cell culture in efforts to develop *in vitro* bioassays to identify compounds that modulate huHDGFh activity or huHDGFh dependent signaling. Accordingly, another aspect of the present invention contemplates a method of identifying compounds efficacious in the treatment of at least one previously described disease or pathology associated with aberrant pre-optic

regulatory factor-2 activity. A non-limiting example of such a method comprises:

a) generating an huHDGFh transgenic non-human animal which is, as compared to a wild-type animal, pathologically distinct in some detectable or measurable manner from wild-type version of said non-human mammal;

b) exposing said transgenic animal to a compound, and;

c) determining the progression of the pathology in the treated transgenic animal, wherein an arrest, delay, or reversal in disease progression in transgenic animal treated with said compound as compared to the progression of the pathology in an untreated control animals is indicative that the compound is useful for the treatment of said pathology

Another embodiment of the present invention provides a method of identifying compounds capable of inhibiting huHDGFh activity *in vivo* and/or *in vitro* wherein said method comprises:

a) administering an experimental compound to an huHDGFh transgenic non-human animal, or tissues derived therefrom, exhibiting one or more physiological or pathological conditions attributable to the overexpression of an huHDGFh transgene; and

b) observing or assaying said animal and/or animal tissues to detect changes in said physiological or pathological condition or conditions.

Another embodiment of the invention provides a method for identifying compounds capable of overcoming deficiencies in huHDGFh activity *in vivo* or *in vitro* wherein said method comprises:

a) administering an experimental compound to an huHDGFh

transgenic non-human animal, or tissues derived therefrom, exhibiting one or more physiological or pathological conditions attributable to the disruption of the endogenous HDGFH gene; and

b) observing or assaying said animal and/or animal tissues to detect changes in said physiological or pathological condition or conditions.

Various means for determining a compound's ability to modulate human huHDGFh in the body of the transgenic animal are consistent with the invention. Observing the reversal of a pathological condition in the transgenic animal after administering a compound is one such means. Another more preferred means is to assay for markers of huHDGFh activity in the blood of a transgenic animal before and after administering an experimental compound to the animal. The level of skill of an artisan in the relevant arts readily provides the practitioner with numerous methods for assaying physiological changes related to therapeutic modulation of huHDGFh activity.

In all previously described *in vitro* and *in vivo* assays, the experimental compound may be administered when applicable, either superficially, orally, parenterally (e.g. by intravenous infusion or injection) or a combination of injection and infusion (iv), intramuscularly (im), or subcutaneously (sc). A preferred route of compound administration to an animal is iv, while oral administration is most preferred.

Having generally described the invention, the same will be more readily understood by reference to the following

examples, which are provided by way of illustration and are not intended as limiting.

Example 1: Expression and Purification of a huHDGFh Polypeptide in E. coli

The bacterial expression vector pQE60 is used for bacterial expression in this example. (QIAGEN, Inc., Chatsworth, CA). pQE60 encodes ampicillin antibiotic resistance ("Ampr") and contains a bacterial origin of replication ("ori"), an IPTG inducible promoter, a ribosome binding site ("RBS"), six codons encoding histidine residues that allow affinity purification using nickel-nitrilo-tri-acetic acid ("Ni-NTA") affinity resin sold by QIAGEN, Inc., and suitable single restriction enzyme cleavage sites. These elements are arranged such that a DNA fragment encoding a polypeptide can be inserted in such a way as to produce that polypeptide with the six His residues (i.e., a "6 X His tag") covalently linked to the carboxyl terminus of that polypeptide. However, a polypeptide coding sequence can optionally be inserted such that translation of the six His codons is prevented and, therefore, a polypeptide is produced with no 6 X His tag.

The nucleic acid sequence encoding the desired portion of a huHDGFh polypeptide lacking the hydrophobic leader sequence is amplified from a cDNA clone using PCR oligonucleotide primers (based on the sequences presented, e.g., as presented in SEQ ID NO:1), which anneal to the amino terminal encoding DNA sequences of the desired portion of a huHDGFh polypeptide and to sequences in the construct 3' to the cDNA coding sequence. Additional nucleotides

containing restriction sites to facilitate cloning in the pQE60 vector are added to the 5' and 3' sequences, respectively.

For cloning a huHDGFh polypeptide, the 5' and 3' primers have nucleotides corresponding or complementary to a portion of the coding sequence of a huHDGFh, e.g., as presented in SEQ ID NO:1, according to known method steps. One of ordinary skill in the art would appreciate, of course, that the point in a polypeptide coding sequence where the 5' primer begins can be varied to amplify a desired portion of the complete polypeptide shorter or longer than the mature form.

The amplified huHDGFh nucleic acid fragments and the vector pQE60 are digested with appropriate restriction enzymes and the digested DNAs are then ligated together. Insertion of the huHDGFh DNA into the restricted pQE60 vector places a huHDGFh polypeptide coding region including its associated stop codon downstream from the IPTG-inducible promoter and in-frame with an initiating AUG codon. The associated stop codon prevents translation of the six histidine codons downstream of the insertion point.

The ligation mixture is transformed into competent E. coli cells using standard procedures such as those described in Sambrook, *et al.*, 1989; Ausubel, 1987-1998. E. coli strain M15/rep4, containing multiple copies of the plasmid pREP4, which expresses the lac repressor and confers kanamycin resistance ("Kanr"), is used in carrying out the illustrative example described herein. This strain, which is only one of many that are suitable for expressing huHDGFh polypeptide, is available commercially from QIAGEN, Inc.

Transformants are identified by their ability to grow on LB plates in the presence of ampicillin and kanamycin. Plasmid DNA is isolated from resistant colonies and the identity of the cloned DNA confirmed by restriction analysis, PCR and DNA sequencing.

Clones containing the desired constructs are grown overnight ("O/N") in liquid culture in LB media supplemented with both ampicillin (100 µg/ml) and kanamycin (25 µg/ml). The O/N culture is used to inoculate a large culture, at a dilution of approximately 1:25 to 1:250. The cells are grown to an optical density at 600 nm ("OD600") of between 0.4 and 0.6. Isopropyl-b-D-thiogalactopyranoside ("IPTG") is then added to a final concentration of 1 mM to induce transcription from the lac repressor sensitive promoter, by inactivating the lacI repressor. Cells subsequently are incubated further for 3 to 4 hours. Cells then are harvested by centrifugation.

The cells are then stirred for 3-4 hours at 4°C in 6M guanidine-HCl, pH8. The cell debris is removed by centrifugation, and the supernatant containing the huHDGFh is dialyzed against 50 mM Na-acetate buffer pH6, supplemented with 200 mM NaCl. Alternatively, a polypeptide can be successfully refolded by dialyzing it against 500 mM NaCl, 20% glycerol, 25 mM Tris/HCl pH7.4, containing protease inhibitors.

If insoluble protein is generated, the protein is made soluble according to known method steps. After renaturation the polypeptide is purified by ion exchange, hydrophobic interaction and size exclusion chromatography. Alternatively, an affinity chromatography step such as an

antibody column is used to obtain pure huHDGFh polypeptide. The purified polypeptide is stored at 4°C or frozen at -40°C to -120°C.

Example 2: Cloning and Expression of a huHDGFh Polypeptide in a Baculovirus Expression System

In this illustrative example, the plasmid shuttle vector pA2 GP is used to insert the cloned DNA encoding the mature polypeptide into a baculovirus to express a huHDGFh polypeptide, using a baculovirus leader and standard methods as described in Summers, *et al., A Manual of Methods for Baculovirus Vectors and Insect Cell Culture Procedures*, Texas Agricultural Experimental Station Bulletin No. 1555 (1987). This expression vector contains the strong polyhedrin promoter of the Autographa californica nuclear polyhedrosis virus (AcMNPV) followed by the secretory signal peptide (leader) of the baculovirus gp67 polypeptide and convenient restriction sites such as BamHI, Xba I, and Asp718. The polyadenylation site of the simian virus 40 ("SV40") is used for efficient polyadenylation. For easy selection of recombinant virus, the plasmid contains the beta-galactosidase gene from E. coli under control of a weak Drosophila promoter in the same orientation, followed by the polyadenylation signal of the polyhedrin gene. The inserted genes are flanked on both sides by viral sequences for cell-mediated homologous recombination with wild-type viral DNA to generate viable virus that expresses the cloned polynucleotide.

Other baculovirus vectors are used in place of the vector above, such as pAc373, pVL941 and pAcIM1, as one

skilled in the art would readily appreciate, as long as the construct provides appropriately located signals for transcription, translation, secretion and the like, including a signal peptide and an in-frame AUG as required. Such vectors are described, for instance, in Luckow, et al., Virology 170:31-39.

The cDNA sequence encoding the mature huHDGFh polypeptide in a clone, lacking the AUG initiation codon and the naturally associated nucleotide binding site, is amplified using PCR oligonucleotide primers corresponding to the 5' and 3' sequences of the gene. Non-limiting examples include 5' and 3' primers having nucleotides corresponding or complementary to a portion of the coding sequence of a huHDGFh polypeptide, e.g., as presented in SEQ ID NO:1, according to known method steps.

The amplified fragment is isolated from a 1% agarose gel using a commercially available kit (e.g., "Geneclean," BIO 101 Inc., La Jolla, CA). The fragment then is then digested with the appropriate restriction enzyme and again is purified on a 1% agarose gel. This fragment is designated herein "F1".

The plasmid is digested with the corresponding restriction enzymes and optionally, can be dephosphorylated using calf intestinal phosphatase, using routine procedures known in the art. The DNA is then isolated from a 1% agarose gel using a commercially available kit ("Geneclean" BIO 101 Inc., La Jolla, CA). This vector DNA is designated herein "V1".

Fragment F1 and the dephosphorylated plasmid V1 are ligated together with T4 DNA ligase. E. coli HB101 or other

suitable E. coli hosts such as XL-1 Blue (Stratagene Cloning Systems, La Jolla, CA) cells are transformed with the ligation mixture and spread on culture plates. Bacteria are identified that contain the plasmid with the human huHDGFh gene using the PCR method, in which one of the primers that is used to amplify the gene and the second primer is from well within the vector so that only those bacterial colonies containing the huHDGFh gene fragment will show amplification of the DNA. The sequence of the cloned fragment is confirmed by DNA sequencing. This plasmid is designated herein pBac huHDGFh .

Five μg of the plasmid pBachuHDGFh is co-transfected with 1.0 μg of a commercially available linearized baculovirus DNA ("BaculoGold™ baculovirus DNA", Pharmingen, San Diego, CA), using the lipofection method described by Felgner, et al., Proc. Natl. Acad. Sci. USA 84:7413-7417 (1987). 1 μg of BaculoGold™ virus DNA and 5 μg of the plasmid pBac huHDGFh are mixed in a sterile well of a microtiter plate containing 50 μl of serum-free Grace's medium (Life Technologies, Inc., Rockville, MD). Afterwards, 10 μl Lipofectin plus 90 μl Grace's medium are added, mixed and incubated for 15 minutes at room temperature. Then the transfection mixture is added drop-wise to Sf9 insect cells (ATCC CRL 1711) seeded in a 35 mm tissue culture plate with 1 ml Grace's medium without serum. The plate is rocked back and forth to mix the newly added solution. The plate is then incubated for 5 hours at 27°C. After 5 hours the transfection solution is removed from the plate and 1 ml of Grace's insect medium supplemented with

10% fetal calf serum is added. The plate is put back into an incubator and cultivation is continued at 27°C for four days.

After four days the supernatant is collected and a plaque assay is performed, according to known methods. An agarose gel with "Blue Gal" (Life Technologies, Inc., Rockville, MD) is used to allow easy identification and isolation of gal-expressing clones, which produce blue-stained plaques. (A detailed description of a "plaque assay" of this type can also be found in the user's guide for insect cell culture and baculovirology distributed by Life Technologies, Inc., Rockville, MD, page 9-10). After appropriate incubation, blue stained plaques are picked with a micropipettor tip (e.g., Eppendorf). The agar containing the recombinant viruses is then resuspended in a microcentrifuge tube containing 200 µl of Grace's medium and the suspension containing the recombinant baculovirus is used to infect Sf9 cells seeded in 35 mm dishes. Four days later the supernatants of these culture dishes are harvested and then they are stored at 4°C. The recombinant virus is called V-huHDGFh.

To verify the expression of the huHDGFh gene, Sf9 cells are grown in Grace's medium supplemented with 10% heat-inactivated FBS. The cells are infected with the recombinant baculovirus V-huHDGFh at a multiplicity of infection ("MOI") of about 2. Six hours later the medium is removed and is replaced with SF900 II medium minus methionine and cysteine (available, e.g., from Life Technologies, Inc., Rockville, MD). If radiolabeled polypeptides are desired, 42 hours later, 5 mCi of 35S-

methionine and 5 mCi 35S-cysteine (available from Amersham) are added. The cells are further incubated for 16 hours and then they are harvested by centrifugation. The polypeptides in the supernatant as well as the intracellular polypeptides are analyzed by SDS-PAGE followed by autoradiography (if radiolabeled). Microsequencing of the amino acid sequence of the amino terminus of purified polypeptide can be used to determine the amino terminal sequence of the mature polypeptide and thus the cleavage point and length of the secretory signal peptide.

Example 3: Cloning and Expression of huHDGFh in Mammalian Cells

A typical mammalian expression vector contains at least one promoter element, which mediates the initiation of transcription of mRNA, the polypeptide coding sequence, and signals required for the termination of transcription and polyadenylation of the transcript. Additional elements include enhancers, Kozak sequences and intervening sequences flanked by donor and acceptor sites for RNA splicing. Highly efficient transcription can be achieved with the early and late promoters from SV40, the long terminal repeats (LTRS) from Retroviruses, e.g., RSV, HTLVI, HIVI and the early promoter of the cytomegalovirus (CMV). However, cellular elements can also be used (e.g., the human actin promoter). Suitable expression vectors for use in practicing the present invention include, for example, vectors such as pIRES1neo, pRetro-Off, pRetro-On, PLXSN, or pLNCX (Clonetech Labs, Palo Alto, CA), pcDNA3.1 (+/-), pcDNA/Zeo (+/-) or pcDNA3.1/Hygro (+/-) (Invitrogen), PSVL

and PMSG (Pharmacia, Uppsala, Sweden), pRSVcat (ATCC 37152), pSV2dhfr (ATCC 37146) and pBC12MI (ATCC 67109). Mammalian host cells that could be used include human Hela 293, H9 and Jurkat cells, mouse NIH3T3 and C127 cells, Cos 1, Cos 7 and CV 1, quail QC1-3 cells, mouse L cells and Chinese hamster ovary (CHO) cells.

Alternatively, the gene can be expressed in stable cell lines that contain the gene integrated into a chromosome. The co-transfection with a selectable marker such as dhfr, gpt, neomycin, or hygromycin allows the identification and isolation of the transfected cells.

The transfected gene can also be amplified to express large amounts of the encoded polypeptide. The DHFR (dihydrofolate reductase) marker is useful to develop cell lines that carry several hundred or even several thousand copies of the gene of interest. Another useful selection marker is the enzyme glutamine synthase (GS) (Murphy, *et al.*, *Biochem. J.* 227:277-279 (1991); Bebbington, *et al.*, *Bio/Technology* 10:169-175 (1992)). Using these markers, the mammalian cells are grown in selective medium and the cells with the highest resistance are selected. These cell lines contain the amplified gene(s) integrated into a chromosome. Chinese hamster ovary (CHO) and NSO cells are often used for the production of polypeptides.

The expression vectors pC1 and pC4 contain the strong promoter (LTR) of the Rous Sarcoma Virus (Cullen, *et al.*, *Molec. Cell. Biol.* 5:438-447 (1985)) plus a fragment of the CMV-enhancer (Boshart, *et al.*, *Cell* 41:521-530 (1985)). Multiple cloning sites, e.g., with the restriction enzyme cleavage sites BamHI, XbaI and Asp718, facilitate the

cloning of the gene of interest. The vectors contain in addition the 3' intron, the polyadenylation and termination signal of the rat preproinsulin gene.

Example 3(a): Cloning and Expression in COS Cells

The expression plasmid, phuHDGFh HA, is made by cloning a cDNA encoding huHDGFh into the expression vector pcDNAI/Amp or pcDNAIII (which can be obtained from Invitrogen, Inc.).

The expression vector pcDNAI/amp contains: (1) an E. coli origin of replication effective for propagation in E. coli and other prokaryotic cells; (2) an ampicillin resistance gene for selection of plasmid-containing prokaryotic cells; (3) an SV40 origin of replication for propagation in eucaryotic cells; (4) a CMV promoter, a polylinker, an SV40 intron; (5) several codons encoding a hemagglutinin fragment (i.e., an "HA" tag to facilitate purification) or HIS tag (see, e.g, Ausubel, supra) followed by a termination codon and polyadenylation signal arranged so that a cDNA can be conveniently placed under expression control of the CMV promoter and operably linked to the SV40 intron and the polyadenylation signal by means of restriction sites in the polylinker. The HA tag corresponds to an epitope derived from the influenza hemagglutinin polypeptide described by Wilson, *et al.*, *Cell* 37:767-778 (1984). The fusion of the HA tag to the target polypeptide allows easy detection and recovery of the recombinant polypeptide with an antibody that recognizes the HA epitope. pcDNAIII contains, in addition, the selectable neomycin marker.

A DNA fragment encoding the huHDGFh is cloned into the polylinker region of the vector so that recombinant polypeptide expression is directed by the CMV promoter. The plasmid construction strategy is as follows. The huHDGFh cDNA of a clone is amplified using primers that contain convenient restriction sites, much as described above for construction of vectors for expression of huHDGFh in E. coli. Non-limiting examples of suitable primers include those based on the coding sequence presented in SEQ ID NO:1, as they encode huHDGFh polypeptides as described herein.

The PCR amplified DNA fragment and the vector, pcDNAI/Amp, are digested with suitable restriction enzyme(s) and then ligated. The ligation mixture is transformed into E. coli strain SURE (available from Stratagene Cloning Systems, 11099 North Torrey Pines Road, La Jolla, CA 92037), and the transformed culture is plated on ampicillin media plates which then are incubated to allow growth of ampicillin resistant colonies. Plasmid DNA is isolated from resistant colonies and examined by restriction analysis or other means for the presence of the huHDGFh-encoding fragment.

For expression of recombinant huHDGFh, COS cells are transfected with an expression vector, as described above, using DEAE-DEXTRAN, as described, for instance, in Sambrook, et al., Molecular Cloning: a Laboratory Manual, Cold Spring Laboratory Press, Cold Spring Harbor, New York (1989). Cells are incubated under conditions for expression of huHDGFh by the vector.

Expression of the huHDGFh-HA fusion polypeptide is detected by radiolabeling and immunoprecipitation, using

methods described in, for example Harlow, *et al.*, *Antibodies:* A Laboratory Manual, 2nd Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York (1988). To this end, two days after transfection, the cells are labeled by incubation in media containing 35S-cysteine for 8 hours. The cells and the media are collected, and the cells are washed and lysed with detergent-containing RIPA buffer: 150 mM NaCl, 1% NP-40, 0.1% SDS, 0.5% DOC, 50 mM TRIS, pH 7.5, as described by Wilson, et al. cited above. Proteins are precipitated from the cell lysate and from the culture media using an HA-specific monoclonal antibody. The precipitated polypeptides then are analyzed by SDS-PAGE and autoradiography. An expression product of the expected size is seen in the cell lysate, which is not seen in negative controls.

Example 3(b): Cloning and Expression in CHO Cells

The vector pC4 is used for the expression of huHDGFh polypeptide. Plasmid pC4 is a derivative of the plasmid pSV2-dhfr (ATCC Accession No. 37146). The plasmid contains the mouse DHFR gene under control of the SV40 early promoter. Chinese hamster ovary- or other cells lacking dihydrofolate activity that are transfected with these plasmids can be selected by growing the cells in a selective medium (alpha minus MEM, Life Technologies) supplemented with the chemotherapeutic agent methotrexate. The amplification of the DHFR genes in cells resistant to methotrexate (MTX) has been well documented (see, e.g., F. W. Alt, *et al.*, *J. Biol. Chem.* 253:1357-1370 (1978); J. L. Hamlin and C. Ma, *Biochem.* et Biophys. Acta 1097:107-143

(1990); and M. J. Page and M. A. Sydenham, *Biotechnology* 9:64-68 (1991)). Cells grown in increasing concentrations of MTX develop resistance to the drug by overproducing the target enzyme, DHFR, as a result of amplification of the DHFR gene. If a second gene is linked to the DHFR gene, it is usually co-amplified and over-expressed. It is known in the art that this approach can be used to develop cell lines carrying more than 1,000 copies of the amplified gene(s). Subsequently, when the methotrexate is withdrawn, cell lines are obtained which contain the amplified gene integrated into one or more chromosome(s) of the host cell.

Plasmid pC4 contains for expressing the gene of interest the strong promoter of the long terminal repeat (LTR) of the Rous Sarcoma Virus (Cullen, *et al.*, *Molec. Cell. Biol.* 5:438-447 (1985)) plus a fragment isolated from the enhancer of the immediate early gene of human cytomegalovirus (CMV) (Boshart, *et al.*, *Cell* 41:521-530 (1985)). Downstream of the promoter are BamHI, XbaI, and Asp718 restriction enzyme cleavage sites that allow integration of the genes. Behind these cloning sites the plasmid contains the 3' intron and polyadenylation site of the rat preproinsulin gene. Other high efficiency promoters can also be used for the expression, e.g., the human b-actin promoter, the SV40 early or late promoters or the long terminal repeats from other retroviruses, e.g., HIV and HTLVI. Clontech's Tet-Off and Tet-On gene expression systems and similar systems can be used to express the huHDGFh in a regulated way in mammalian cells (M. Gossen, and H. Bujard, *Proc. Natl. Acad. Sci.* USA 89: 5547-5551 (1992)). For the polyadenylation of the mRNA other signals,

e.g., from the human growth hormone or globin genes can be used as well. Stable cell lines carrying a gene of interest integrated into the chromosomes can also be selected upon co-transfection with a selectable marker such as gpt, G418 or hygromycin. It is advantageous to use more than one selectable marker in the beginning, e.g., G418 plus methotrexate.

The plasmid pC4 is digested with restriction enzymes and then dephosphorylated using calf intestinal phosphatase by procedures known in the art. The vector is then isolated from a 1% agarose gel.

The DNA sequence encoding the complete huHDGFh polypeptide is amplified using PCR oligonucleotide primers corresponding to the 5' and 3' sequences of the gene. Non-limiting examples include 5' and 3' primers having nucleotides corresponding or complementary to a portion of the coding sequence of a huHDGFh, e.g., as presented in SEQ ID NO:1, according to known method steps.

The amplified fragment is digested with suitable endonucleases and then purified again on a 1% agarose gel. The isolated fragment and the dephosphorylated vector are then ligated with T4 DNA ligase. E. coli HB101 or XL-1 Blue cells are then transformed and bacteria are identified that contain the fragment inserted into plasmid pC4 using, for instance, restriction enzyme analysis.

Chinese hamster ovary (CHO) cells lacking an active DHFR gene are used for transfection. 5 μg of the expression plasmid pC4 is cotransfected with 0.5 μg of the plasmid pSV2-neo using lipofectin. The plasmid pSV2neo contains a dominant selectable marker, the neo gene from Tn5 encoding

an enzyme that confers resistance to a group of antibiotics including G418. The cells are seeded in alpha minus MEM supplemented with 1 μg/ml G418. After 2 days, the cells are trypsinized and seeded in hybridoma cloning plates (Greiner, Germany) in alpha minus MEM supplemented with 10, 25, or 50 ng/ml of methotrexate plus 1 μg/ml G418. After about 10-14 days single clones are trypsinized and then seeded in 6-well petri dishes or 10 ml flasks using different concentrations of methotrexate (50 nM, 100 nM, 200 nM, 400 nM, 800 nM). Clones growing at the highest concentrations of methotrexate are then transferred to new 6-well plates containing even higher concentrations of methotrexate (1 mM, 2 mM, 5 mM, 10 mM, 20 mM). The same procedure is repeated until clones are obtained which grow at a concentration of 100 - 200 mM. Expression of the desired gene product is analyzed, for instance, by SDS-PAGE and Western blot or by reverse phase HPLC analysis.

Example 4: Tissue Distribution of huHDGFh mRNA Expression

Northern blot analysis is carried out to examine huHDGFh gene expression in human tissues, using methods described by, among others, Sambrook, *et al.*, cited above. A cDNA probe containing the entire nucleotide sequence of a huHDGFh polypeptide (SEQ ID NO:1) is labeled with ^{32}P using the Rediprime™ DNA labeling system (Amersham Life Science), according to the manufacturer's instructions. After labeling, the probe is purified using a CHROMA SPIN-100™ column (Clontech Laboratories, Inc.), according to the manufacturer's protocol number PT1200-1. The purified and

labeled probe is used to examine various human tissues for huHDGFh mRNA.

Multiple Tissue Northern (MTN) blots containing various human tissues (H) or human immune system tissues (IM) are obtained from Clontech and are examined with the labeled probe using ExpressHyb hybridization solution (Clontech) according to manufacturer's protocol number PT1190-1. Following hybridization and washing, the blots are mounted and exposed to film at -70°C overnight, and films developed according to standard procedures. The results show huHDGFh polypeptides to be selectively expressed in other tissues.

It will be clear that the invention can be practiced otherwise than as particularly described in the foregoing description and examples.

Numerous modifications and variations of the present invention are possible in light of the above teachings and, therefore, are within the scope of the appended claims.

What is claimed is:

1. An isolated nucleic acid, comprising a polynucleotide encoding at least 90-100% of the contiguous amino acids of a protein sequence selected from SEQ ID NO:2.

2. An isolated nucleic acid, comprising at least one huHDGFh polynucleotide encoding at least 90-100% of the contiguous amino acids of a protein sequence selected from SEQ ID NO:2, further comprising at least one mutation corresponding to at least one substitution, insertion or deletion selected from the group consisting of 9F, 11P, 15I, 29V, 31D, 32V, 32I, 33A, 34D, 41T, 41P, 44L, 62I, 66S, 66D, 67K, 68N, 68C, 70E, 72Y, 75P, 89D, 89Q, 93K, 93H, 94A, 95S, 96Y, 97S, 98S, 98A, 99Q, 99Y, 100P, 101P, 103T, 103S, 104K, 104S, and 105S of SEQ ID NO:2.

3. An isolated nucleic acid, comprising at least one huHDGFh polynucleotide comprising or complementary to at least 90-100% of the contiguous nucleotides of SEQ ID NO:1.

4. A composition, comprising at least one isolated nucleic acid according to any of claims 1-3 and a carrier or diluent.

5. A recombinant vector, comprising at least one nucleic acid according to any of claims 1-3.

6. A host cell comprising at least one recombinant vector according to claim 5.

7. A method for producing at least one huHDGFh polypeptide, comprising culturing a host cell according to claim 6 under conditions that the at least one huHDGFh polypeptide is expressed in detectable or recoverable amounts.

8. A transgenic or chimeric non-human animal, comprising at least one isolated nucleic acid according to any of claims 1-3.

9. An isolated polypeptide, comprising a huHDGFh polypeptide comprising at least 90-100% of the contiguous amino acids of at least one amino acid sequence of SEQ ID NO:2.

10. An isolated polypeptide according to claim 9, further comprising at least one mutation corresponding to at least one substitution, insertion or deletion selected from the group consisting of 9F, 11P, 15I, 29V, 31D, 32V, 32I, 33A, 34D, 41T, 41P, 44L, 62I, 66S, 66D, 67K, 68N, 68C, 70E, 72Y, 75P, 89D, 89Q, 93K, 93H, 94A, 95S, 96Y, 97S, 98S, 98A, 99Q, 99Y, 100P, 101P, 103T, 103S, 104K, 104S, and 105S of SEQ ID NO:2.

11. An isolated polypeptide comprising at least one polypeptide comprising at least 90-100% of the contiguous amino acids of at least one extracellular, intracellular, transmembrane or active domain of SEQ ID NO:2.

12. A composition, comprising at least one isolated polypeptide according to any of claims 8-11 and a carrier or diluent.

13. An isolated nucleic acid probe, fragment, or primer, comprising a huHDGFh polynucleotide comprising a sequence corresponding or complementary to at least 10 nucleotides of SEQ ID NO:1.

14. An isolated nucleic acid, comprising a nucleic acid that hybridizes under stringent conditions to a nucleic acid according to claim 13.

15. An antibody or at least one fragment thereof that binds an epitope specific to at least one huHDGFh polypeptide according to any of claims 8-11.

16. A host cell, expressing at least one antibody or at least one fragment thereof according to claim 15.

17. A method for producing at least one antibody, comprising culturing a host cell according to claim 16.

18. A method for identifying compounds that bind at least one huHDGFh polypeptide, comprising

(a) admixing at least one isolated huHDGFh polypeptide according to any of claims 8-11 with at least one test compound or composition; and

(b) detecting at least one binding interaction between said at least one huHDGFh polypeptide and the test compound or composition.

19. A compound or composition detected by method according to claim 18.

20. Any invention described herein.

SEQUENCE LISTING

<110> Na, Songqing

<120> huHDGFh NUCLEIC ACIDS, POLYPEPTIDES, METHODS AND USES THEREOF

<130> X-12275

<140>
<141>

<160> 2

<170> PatentIn Ver. 2.0

<210> 1
<211> 609
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (1)..(609)

<400> 1

```
atg gcg cgt ccg cgg ccc cgc gag tac aaa gcg ggc gac ctg gtc ttc   48
Met Ala Arg Pro Arg Pro Arg Glu Tyr Lys Ala Gly Asp Leu Val Phe
  1               5                  10                  15

gcc aag atg aag ggc tac ccg cac tgg ccg gcc cgg att gat gaa ctc   96
Ala Lys Met Lys Gly Tyr Pro His Trp Pro Ala Arg Ile Asp Glu Leu
             20                  25                  30

cca gag ggc gct gtg aag cct cca gca aac aag tat cct atc ttc ttt   144
Pro Glu Gly Ala Val Lys Pro Pro Ala Asn Lys Tyr Pro Ile Phe Phe
         35                  40                  45

ttt ggc acc cat gaa act gca ttt cta ggt ccc aaa gac ctt ttt cca   192
Phe Gly Thr His Glu Thr Ala Phe Leu Gly Pro Lys Asp Leu Phe Pro
     50                  55                  60

tat aag gag tac aaa gac aag ttt gga aag tca aac aaa cgg aaa gga   240
Tyr Lys Glu Tyr Lys Asp Lys Phe Gly Lys Ser Asn Lys Arg Lys Gly
 65                  70                  75                  80

ttt aac gaa gga ttg tgg gaa ata gaa aat aac cca gga gta aag ttt   288
Phe Asn Glu Gly Leu Trp Glu Ile Glu Asn Asn Pro Gly Val Lys Phe
```

85 90 95

act ggc tac cag gca att cag caa cag agc tct tca gaa act gag gga 336
Thr Gly Tyr Gln Ala Ile Gln Gln Gln Ser Ser Ser Glu Thr Glu Gly
100 105 110

gaa ggt gga aat act gca gat gca agc agt gag gaa gaa ggt gat aga 384
Glu Gly Gly Asn Thr Ala Asp Ala Ser Ser Glu Glu Glu Gly Asp Arg
115 120 125

gta gaa gaa gat gga aaa ggc aaa aga aag aat gaa aaa gca ggc tca 432
Val Glu Glu Asp Gly Lys Gly Lys Arg Lys Asn Glu Lys Ala Gly Ser
130 135 140

aaa cgg aaa aag tca tat act tca aag aaa tcc tct aaa cag tcc cgg 480
Lys Arg Lys Lys Ser Tyr Thr Ser Lys Lys Ser Ser Lys Gln Ser Arg
145 150 155 160

aaa tct cca gga gat gaa gat gac aaa gac tgc aaa gaa gag gaa aac 528
Lys Ser Pro Gly Asp Glu Asp Asp Lys Asp Cys Lys Glu Glu Glu Asn
165 170 175

aaa agc agc tct gag ggt gga gat gcg ggc aac gac aca aga aac aca 576
Lys Ser Ser Ser Glu Gly Gly Asp Ala Gly Asn Asp Thr Arg Asn Thr
180 185 190

act tca gac ttg cag aaa acc agt gaa ggg acc 609
Thr Ser Asp Leu Gln Lys Thr Ser Glu Gly Thr
195 200

<210> 2
<211> 203
<212> PRT
<213> Homo sapiens

<400> 2
Met Ala Arg Pro Arg Pro Arg Glu Tyr Lys Ala Gly Asp Leu Val Phe
1 5 10 15

Ala Lys Met Lys Gly Tyr Pro His Trp Pro Ala Arg Ile Asp Glu Leu
20 25 30

Pro Glu Gly Ala Val Lys Pro Pro Ala Asn Lys Tyr Pro Ile Phe Phe
35 40 45

Phe Gly Thr His Glu Thr Ala Phe Leu Gly Pro Lys Asp Leu Phe Pro
50 55 60

Tyr Lys Glu Tyr Lys Asp Lys Phe Gly Lys Ser Asn Lys Arg Lys Gly
65 70 75 80

Phe Asn Glu Gly Leu Trp Glu Ile Glu Asn Asn Pro Gly Val Lys Phe
85 90 95

Thr Gly Tyr Gln Ala Ile Gln Gln Gln Ser Ser Ser Glu Thr Glu Gly
100 105 110

Glu Gly Gly Asn Thr Ala Asp Ala Ser Ser Glu Glu Glu Gly Asp Arg
115 120 125

Val Glu Glu Asp Gly Lys Gly Lys Arg Lys Asn Glu Lys Ala Gly Ser
130 135 140

Lys Arg Lys Lys Ser Tyr Thr Ser Lys Lys Ser Ser Lys Gln Ser Arg
145 150 155 160

Lys Ser Pro Gly Asp Glu Asp Asp Lys Asp Cys Lys Glu Glu Glu Asn
165 170 175

Lys Ser Ser Ser Glu Gly Gly Asp Ala Gly Asn Asp Thr Arg Asn Thr
180 185 190

Thr Ser Asp Leu Gln Lys Thr Ser Glu Gly Thr
195 200